DOCUMENT RESUME

ED 236 849 EC 160 867

AUTHOR Skiba, Russell J.
TITLE The Relationship between Classroom Management Strategies and Student Misbehaviors.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
REPORT NO IRLD-RR-133
PUB DATE Jul 83
CONTRACT 300-80-0622
NOTE 65p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Behavior Problems; *Classroom Techniques; *Teacher Behavior
IDENTIFIERS *Classroom Management Observation Scale

ABSTRACT

Because research has determined that specific management techniques can have an effect on the classroom behavior of students, an observational rating scale was developed to assess the type of management techniques six elementary teachers in a program for behaviorally disordered children used to control behavior. Correlational analyses were used to determine the relationships among teacher behaviors, and between teacher behavior and measures of student misbehavior. Results clearly showed the importance of direct management variables over indirect structuring variables, and in particular the importance of managing inappropriate behavior. Appended materials include operational definitions for the classroom management observation scale (CMOS) and information on the technical characteristics of the CMOS. (Author/CL)

University of Minnesota

Research Report No. 133

THE RELATIONSHIP BETWEEN CLASSROOM MANAGEMENT STRATEGIES AND STUDENT MISBEHAVIORS

Russell J. Skiba

IRLD

Institute for Research on Learning Disabilities

IRLD

Director: James E. Ysseldyke

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Office of Special Education.

Research Report No. 133

THE RELATIONSHIP BETWEEN CLASSROOM MANAGEMENT STRATEGIES AND STUDENT MISBEHAVIORS

Russell J. Skiba

Institute for Research on Learning Disabilities

University of Minnesota

July, 1983

Abstract

Research has determined that specific management techniques can have an effect on the classroom behavior of students. An observational rating scale was developed, based on these findings, to assess the type of management techniques used to control behavior. Subjects were six teachers in a program for behaviorally disordered children. Correlational analyses were used to determine the relationships among teacher behaviors, and between teacher behavior and measures of student misbehavior. Implications for educational research and practice are discussed.

The Relationship Between Classroom Management Strategies and Student Misbehaviors

The importance of classroom misbehavior, and teacher management of that behavior, cannot be underestimated in either regular or special education. During the 1978-79 school year, one of every 20 public school teachers in urban areas reported having been physically attacked on school property (National Institute of Education, 1978), a factor some have suggested might contribute to a higher turnover rate among teachers (McGuire, 1979). Student misbehavior also has been shown to be related negatively to academic achievement. Research has shown that a strong relationship obtains between measures of student attention and academic achievement (Hoge & Luce, 1979), and that the relationship is strongest for low-achieving students (Soli & Devine, 1976). Inappropriate classroom behavior also may play a key role in referral to special education placement (cf. Algozzine, Ysseldyke, & Christenson, 1983; Mirkin, Marston, & Deno, in press).

In recent years, research has begun to focus on the observation of specific teacher behaviors and the relation of these to student outcomes. Various research programs have identified a number of teaching variables important in the control and direction of student behavior in the classroom. These variables can be grouped into two types: immediate and long-term. Immediate variables are those that have been observed to have consistent effects on the observed behavior of students, in terms of fewer disruptive behaviors, greater attention to task, and so on. Long-term variables are those that have been correlated with student outcomes that are less directly observable, such as grades and scores on achievement tests. An extensive review

of these variables is provided by Skiba (1983). A brief summary of the variables is provided here.

Immediate Variables

The positive effects of teacher attention in increasing appropriate behavior and decreasing inappropriate behavior have been demonstrated in a number of studies (e.g., Cooper, Thomson, & Baer, 1970; Hall, Lund, & Jackson, 1968; Madsen, Becker, & Thomas, 1968). O'Leary and O'Leary (1977) outlined three characteristics necessary for teacher attention to function as a reinforcer: it must be contingent upon behavior, it must specify the desired behavior, and it must be delivered by the teacher in a sincere and credible manner. Inconsistent results regarding the use of praise as a reinforcer (Good & Grouws, 1977), however, have led some to suggest that praise is not always intended as a reinforcer, since it is often not delivered in accordance with these specifications (Brophy, 1981).

The use of token economies also has been found to be effective in decreasing disruptive behavior and shaping appropriate behavior (Kazdin, 1977; O'Leary & Drabman, 1971). Reinforcers that have been found to be effective in such systems include free-time (Conch & Clement, 1981), sports activities (Hansen, McLaughlin, Hansaker, & Young, 1981), and video-games (Robinson, Newby, & Ganzell, 1981).

Many writers have suggested that when behavior problems are severe, or are being reinforced by peer attention or other environmental contingencies, some form of direct intervention may be necessary, in conjunction with a positive program, to control the behavior (Jones & Miller, 1974; Mattos, Mattson, Walker, & Buckley,

1969). A number of interventions have been shown to successfully reduce classroom misbehavior. Of these, time out from positive reinforcement has been found to be successful most often (cf. Powell & Powell, 1982). Other successful interventions have included contingent after-school time (Swanson, 1979), loss of tokens or "response-cost" (Pace & Foreman, 1982), and the use of reprimands delivered in a soft tone of voice (O'Leary, Kaufman, Kass, & Drabman, 1970).

The effectiveness of classroom management also may be mediated by the degree or type of cognitive structure present, i.e., the use of classroom rules. Herman and Tramontana (1971) reported that token reinforcement failed to decrease high rates of misbehavior unless used in conjunction with specific instructions about appropriate and inappropriate behavior. Evertson and Emmer (1982) reported that rules were most effective when presented clearly and enforced consistently.

The work of Kounin (1970) suggests that more effective classroom managers evidence a higher degree of awareness (i.e., "withitness"). Teachers who more often ask students about their progress on assigned work and alert off-task students that they are liable to have their work checked have been shown to have significantly fewer behavior problems in their classrooms (Borg & Ascione, 1982). Evertson and Emmer (1982) reported that the most highly significant difference between effective and ineffective classroom managers was on the variable "effectively monitors student progress and completion of assignments."

Long-Term Variables

One of the more significant findings of observational teacher effectiveness research has been that amount of student learning is directly proportional to amount of time actively engaged in academic learning. In the Beginning Teacher Evaluation Study (Fisher, Berliner, Filby, Marliave, Cahen, & Dishaw, 1980) both the amount of time allocated to student instruction by the teacher, and the proportion of time the student was actively engaged in learning correlated positively with student performance on achievement tests.

The characteristics of the instruction that students receive also have been explored in relation to academic outcomes. Teacher variables that have been shown to be important in predicting academic outcome have included teacher feedback and corrections of student responding (Anderson, Evertson, & Brophy, 1979; Brophy & Evertson, 1976), highly structured teacher-directed academic activities (Stallings, 1975; Stevens & Rosenshine, 1981), smooth transitions between subjects during lessons (Anderson et al., 1979; Kounin, 1970), and a rapid lesson pace (Good, Grouws, & Beckerman, 1978; Stallings, 1975).

Because of the different theoretical and methodological backgrounds of the researchers investigating these concepts, however, rarely are more than a few of these variables studied simultaneously. Further, although some have asserted that teaching behaviors that structure the overall learning environment are more important than direct behavior management in maintaining classroom discipline (Kounin, 1970), few studies have directly compared the effects of the

two types of teacher behavior. In addition, with the exception of reinforcement and punishment, few of the variables identified as important in classroom management have been studied in settings devoted to special populations. It seems likely that such environments, smaller classrooms composed of children with more intense behavior or learning problems, would require different teacher behaviors for success than would larger, regular classroom settings.

The purpose of the current study was to observe and compare the effectiveness of various teacher behaviors in the management of behavior disordered children. Using the findings of observational classroom literature, a rating scale was developed to allow observation of a number of teacher behaviors simultaneously. In addition to comparisons of specific behaviors, aggregation of the scales facilitated comparison of the effectiveness of immediate and long-term variables. The effects of teacher attention and general classroom structure on student behavior also were compared.

Method

Subjects

Subjects were six teachers in a public school program for behavior disordered children in a large upper midwestern city. The majority of the students in the program had been referred for inappropriate behavior in the classroom.

The program for these students was more structured than the regular education program. A school time-out room (with time-out room aides) was available when classroom behavior became too severe, and access to privileges was contingent upon behavior by means of a code

system. In addition, most of the teachers used token economies in their classrooms on a regular basis.

Teachers in the program, however, differed in the extent to which they used these procedures. Although the time to be spent in time out was specified by school regulations, no regulations governed when a student had to be sent to time out. Only 53% of the teachers reported using time out regularly in the control of misbehavior. Similarly, use of the point system varied. Of all the teachers in the program, 67% reported the use of some form of point system; 60% used some form of edible reinforcers, and only 20% used actual tokens as a part of their management system. Thus, there was considerable variability in the implementation of the behavioral procedures within the program. This variability was considered crucial for the purposes of the study, since it allowed observation of a variety of management styles.

All subjects were asked to participate in the study by the program coordinator. Eight teachers originally agreed to participate, but two later asked not to be included. Three of the six participating teachers worked with students in grades 1-3 (median age = 7.5), and three worked with students in grades 4-6 (median age = 10.2). The mean number of years teaching in the program was 3.5; the mean number of years teaching special education was 7.8. The median number of students in participating classrooms was 9.5.

Measures and Observer Training

The Classroom Management Observation Scale (CMOS) was designed to assess the style of teacher management in the classroom. It consists of 11 scales: 10 individual five-point scales and one variable rated

on a two-point scale of present or absent. All were uniformly scaled in a positive direction, that is, a higher rating represented a greater frequency or degree of the behavior being observed; behavioral definitions were provided for both ends of each scale. The scales represented variables that prior research had shown to be correlated with improved student behavior or higher academic achievement. Defining characteristics of the variables, as well as respresentative studies that have investigated these characteristics, are presented in Table 1. Operational definitions actually used by the raters and a sample protocol are included in Appendix A.

Insert Table 1 about here

Four observers were trained in the use of the CMOS and given the opportunity to practice rating two videotapes of classroom situations (only three of these observers later participated in classroom observation; the investigator served as the fourth observer). The first tape showed a teacher described as highly effective by both peers and superiors in a regular classroom situation with three students. After viewing and rating the first tape, observers compared and discussed ratings and came to a consensus regarding the "appropriate" rating on each variable. They then viewed and rated a second tape, in which the same teacher had been instructed to "act as a poor teacher would." These two tapes gave raters the experience of rating at both ends of the scales.

Inter-rater agreements, defined as ratings within one scale point, were calculated for all observers on both tapes. Agreement

coefficients between observers on the first tape ranged from .82 to 1.00, with a mean of .89; observations of the second tape yielded coefficients ranging from .73 to .81, with a mean of .85. Only two coefficients out of 20 failed to exceed .80, the level of agreement usually recommended in the literature (Keller, 1980).

The dependent variable, student misbehavior, was measured by means of a frequency count system devised by Deno (1979). This method combines time sampling and frequency counts; the observer records all instances of a given behavior in a 20-second interval, then counts all instances of the next behavior category for the next 20 seconds, and so on. The original system provides four categories of behavior: noise, out-of-place, physical contact or destruction, and off-task. During the first classroom observation, however, the frequency with which physical contacts occurred was extremely low. Thus, this category was dropped to allow additional observation time for the other three categories. Definitions of the behaviors observed, and a sample protocol, may be found in Appendix B. Observer training for the behavior coding system also was provided by means of videotape. Inter-observer agreement coefficients between two pairs of observers were .67 and .90.

Reliability checks also were conducted for both measurement instruments during classroom observation, as a check on what O'Leary and Kent (1973) refer to as "observer drift." Inter-observer agreement across all CMOS scales ranged from .80 to .82. Analyses of the inter-rater agreement for each CMOS scale indicated that most scales showed good agreement, with the exception of classroom rules

(.62), positive attention (.57), and transitions (.57). All other scales showed within-one point agreements ranging from .89 to 1.00. However, exact agreement among observers was considerably lower, ranging from .20 to .50 across scales, and from .40 to .82 for the 11 scales. Inter-observer agreement on the Classroom Behavior Frequency Counts ranged from .70 to 1.00 (average for total score = .82) when the criterion was the same number of behaviors per interval. When the agreement criterion was that the same behavior was coded during an interval, agreement ranged from .30 to 1.00, with the average for the total score being .67.

Procedures

Classrooms were observed three times during the fall of the school year. The first observation took place in October, six weeks after the beginning of the school year. The second observation occurred three weeks later, and the last four weeks after that, shortly before the December break. Classroom observers were stationed in observation rooms adjoining the classroom. The four observers were rotated in all classrooms, and no observer was in a given classroom more than once.

Each session began with 20 minutes of observation time on the CMOS. During this time, observers were encouraged to refer to the scale definitions frequently to guide their observations, but were instructed not to rate until the full 20 minutes had passed. A five minute block of time was allowed for scoring the teacher on the 11 scales. The observer then recorded student misbehavior on the Classroom Behavior Frequency Count for 20 minutes. Clipboards

equipped with a timing device and earphone signaled the beginning of each new interval to the observer.

Data Analysis

Descriptive data are reported for each of the three observations. Correlations among the dependent variables (student misbehavior) and the independent variables (teacher classroom management ratings) are based on the aggregation of data across all three observations for both dependent and independent variables. Since classroom size ranged from 5 to 10 students during observation periods, all student behavior data are controlled for class size. Although the correlational data are useful in describing relationships among the variables, caution is advised in interpreting or attempting to generalize the results, due to the small sample size. Statistical analyses also were performed to determine the reliability of measurement instruments. The results of these analyses are reported in Appendix C.

Results

Description of Teacher and Student Behaviors

Teacher behavior. The means and standard deviations for the 11 CMOS scale items indicated that, on the average, use of the variables was high in the classrooms observed (see Table 2). These results may be somewhat skewed, however, by the high use of the variables by the teachers in grades 4-6. Breakdown of the variables by primary (grades 1-3) and intermediate (grades 4-6) classrooms showed large differences between the two groups of teachers on all the variables except positive attention and back-up reinforcers (see Table 3).

Insert Tables 2 and 3 about here

Analysis of the data across observations did yield some consistent trends. Use of feedback and corrections, and classroom rules, tended to increase over time; smooth transitions, structure of the lesson, and pacing and enthusiasm tended to decrease over time. Given the standard deviations of the items, and the low exact reliability of the scale, these results must be interpreted with caution.

Student behavior. Actual numbers of student misbehaviors varied widely across classrooms. During the third observation, for instance, the total number of inappropriate behaviors counted in any given classroom (unadjusted for class size) ranged from 0 to 122 over the course of the 20-minute observation period.

Analysis of the number of behaviors emitted per child (see Table 4) and percentage of intervals in which a behavior occurred (see Table 5) yielded somewhat contradictory results: mean number of off-task and total behaviors per child appeared to increase dramatically over time, while the same data measured as a proportion of total time appeared to decrease. In fact, the truth may lie somewhere in between. While frequencies of the observed behaviors did tend to increase over time in all classrooms, the results at Time 3 were somewhat skewed by a dramatic two-fold increase in one of the classrooms. The proportional interval data tend to underestimate the behavior, since the use of proportions imposes a ceiling effect on the

data (Haynes, 1978). Nevertheless, both approaches indicate that, in the classrooms studied, off-task behaviors occurred most frequently, followed by noises and out-of-seat behaviors.

Insert Tables 4 and 5 about here

Relationships among teacher behaviors. Pearson product correlations among the aggregate ratings of teacher behaviors are presented in Table 6. The results indicated that, with the exception of positive attention and back-up reinforcers, the variables were moderately to highly correlated with each other. A number of interesting relationships emerge in viewing the strongest correlations. Teachers who intervened most briefly, consistently, and specifically when dealing with inappropriate behaviors also responded more immediately and were in general more aware of the students in their classrooms. Teachers with more structured lessons also tended to deliver more feedback to their students regarding that lesson. Classrooms in which the teacher exhibited greater awareness of student activity tended to be classrooms that also exhibited smoother transitions, and a higher rate of active academic responding. Finally, strong correlations were obtained between pacing and interventions for inappropriate behavior, immediacy of consequences, teacher awareness, and transitions.

Insert Table 6 about here

Relationships among student behaviors. Correlations among student behaviors were strong and stable over time. Table 7 presents the correlations among observed behaviors, aggregated across sessions.

Insert Table 7 about here

Relationships among teacher and student behaviors. Correlations between ratings on the CMOS and student misbehavior were, in general, quite strong (see Table 8). Although high correlations were obtained regardless of treatment of student behavior data, only correlations based on the proportion of intervals in which misbehavior occurred are presented, since this method yielded greater agreement coefficients among observers. Since higher ratings on the individual scales indicate greater classroom structure, one would predict negative correlations with student misbehavior: the higher the ratings, the lower the frequency of misbehavior. The results reflect such a trend. Only back-up reinforcers and positive attention for appropriate behavior were mildly positively related to inappropriate behavior. Interventions that correlated most highly with a classroom free of student misbehavior were brief, specific, and unemotional; immediate consequences, consistent classroom rules, and a brisk lesson pace also correlated relatively highly with low rates of student misbehavior.

Insert Table 8 about here

Three sets of aggregate scales were formed: the first was based on a distinction between immediate and long-term variables; the second

consisted of a teacher attention scale containing those variables relating to teacher attention to individual students; the third was a general structure scale, composed of those variables relating to teacher behaviors that functioned to structure the classroom environment. Although little difference was found in the correlations between teacher attention and student behavior, and between general structure and student behavior, immediate variables involving management of student behavior correlated more highly, in a negative direction, with misbehavior than did long-term variables.

Correlated vs. effective variables. Although a number of variables correlated negatively with frequency of misbehavior, the high correlations among the teaching variables may indicate that some of the teaching behaviors relate to decreased disruptive behavior only because of their correlation with a more powerful teaching behavior. Put another way, some of the variables may just be descriptions of behaviors that effective classroom managers also happen to engage in, although such behaviors do not, in and of themselves, control classroom behavior.

Although a full analysis of causation and correlation would necessitate experimental manipulation of teacher behavior, some statistical control is possible through partial correlations, where the simple correlation between two variables is controlled for a third variable that may also correlate highly with one or the other variable. Although sample size is, in this case, too small to allow judgments about statistical significance, some trends can be noted by comparing simple correlations between a teacher variable and student

behavior, and correlations between those variables controlled for other teacher behaviors.

In a series of partial correlations, interventions for inappropriate behavior and immediacy of consequences were the only variables unaffected by controlling for other teaching variables. The simple correlations for interventions for inappropriate behavior were around -.95; controlling for other variables yielded partial correlations that ranged from -.60 to -.96, with the majority of correlations around -.90. Immediacy of consequences was more affected, but first order partials still ranged from -.51 to -.97, with the majority of correlations around -.90. Other correlations changed dramatically by partialling out the variance due to other variables: teacher awareness, transitions, pacing and enthusiasm, and classroom rules (variables that had correlated strongly with low rates of misbehavior) evidenced weak negative or even positive correlations with student misbehavior when controlled for interventions for inappropriate behavior or immediacy of consequences. Feedback and corrections, and lesson structure may be mutually dependent; although correlations with student behavior remained healthy when controlled for other variables, when controlled for each other, originally strong negative correlations with misbehavior were reduced to -.39 and -.37, respectively.

Further information also was provided about aggregated variables through partial correlations. While the correlation between the immediate variables and total misbehaviors dropped only moderately, to -.50, when controlled by long-term variables, the partial correlation

between long-term variables and misbehavior was .09 when partialling out the effect of immediate variables. Teacher attention showed a similar stability in comparison to general structure, suggesting that whatever power long-term structuring variables have in controlling classroom behavior arises only in conjunction with immediate management variables involving teacher attention.

Discussion

One of the difficulties with observational classroom research, more common to the broad studies of process-product research than to the narrower focus of behavioral research, might be called the problem of correlated vs. effective variables. While many variables have been shown to relate to student outcomes in purely correlational studies, it seems likely that at least some of these variables acquire their predictive power solely through their association with variables that may play a more direct causal role. Partial correlations among the CMOS variables revealed that only interventions for inappropriate behavior and immediacy of consequences were relatively unaffected by controlling for other behaviors. The same analysis revealed that feedback and lesson structure seemed mutually dependent, although relatively unaffected by other variables. A similar process is doubtless at work in the regression analyses of many process-product studies, and may in part account for the small percentage of variance explained by many process-product variables.

Another point of interest was the strong relationship among the different behavior codes. Noise, out of seat, off task, and their summation all correlated highly with each other, and acted similarly

in relationship with other variables. This may have something to do with the definitions of student behavior. These behavior codes, like many used in educational research, tend to be topographic and structural: the important differences between the operational definitions of 'noise' and 'out-of-seat' concern the parts of the body with which they are emitted. Thompson and Lubinski (1982) argue that behavior is defined more aptly through a functional analysis, that is, in terms of the stimuli and consequences that shape and maintain it. The present results support such an interpretation, since behaviors that differ topographically seem to have the same controlling variables, providing some evidence of functional equivalence.

The results clearly showed, for this sample of classrooms, the importance of direct management variables over indirect structuring variables, and in particular the importance of managing inappropriate behavior. While the small sample size does make generalization to any larger population risky, some observations can be made. The current results tend not to support the hypothesis of Kounin (1970) that indirect structuring variables such as smoothness and momentum (transitions and pacing in the current study) are more important in maintaining classroom discipline than teacher behaviors specifically intended to manage behavior. They do support the experimental findings of Jones and Miller (1974) and others that some form of direct intervention is necessary to control misbehavior, especially for children with more severe behavioral problems, and that such interventions are most effective when delivered immediately, briefly and unemotionally, and with a degree of cognitive structure.

The failure of either social or tangible reinforcement to correlate with decreased disruption in this sample was surprising. One explanation may lie in the fact that positive teacher attention and back-up reinforcers were used with a certain uniformity in all classrooms observed, while interventions for inappropriate behavior varied widely. This suggests that praise and token systems in such classrooms only acquire their power in conjunction with actions designed to control behaviors directly. If children are receiving peer attention (or even teacher attention) for misbehavior, positive reinforcement alone may be insufficient to extinguish the inappropriate behavior.

Previous research has hinted at developmental limitations for some of the teaching variables studied here. Both positive attention and teacher feedback have been demonstrated to be related to student outcomes at some grade levels, but not at others (Stallings, 1975). It is easy to imagine how other variables, such as classroom rules, would also be affected by student developmental level. Thus, the differences in implementation of the behaviors between the primary and intermediate teachers cannot be taken as an index of teacher quality, or the assumption made that, if trained in these variables, the primary classrooms would evidence lower rates of disruption. An alternative hypothesis might suggest that such teaching behaviors are more difficult to carry out with younger children.

Although they must be regarded as preliminary given the size of the sample on which they are based, the current findings suggest that caution may be in order in attempting to generalize the findings of

process-product research to special populations. While most of the variables did correlate with student behavior outcomes, partial correlations revealed that many of the teaching strategies acquired their predictive power only in conjunction with other, more directly correlated variables. Thus, teaching behaviors found effective in regular classroom management may be dependent on their relationship with more direct management strategies in special classroom settings.

Active interventions designed to extinguish classroom misbehavior proved to be the most important dimension of classroom behavior management for this sample. Such results serve to underscore the need for further research into the parameters of punishment-related management strategies, such as time out, reprimands, or response cost. Ethical objections to such techniques may be well intentioned, but will not substitute for actual data, especially when such techniques are already in use in special education programs. While ethical considerations may preclude radical manipulation of punishment-related interventions experimentally, the use of such techniques varies widely enough (even in the present smaller sample) to permit observational comparisons of naturally occurring interventions.

The new observational research methodologies in classroom research do offer a hopeful sign. Static variables, such as the label "emotional disturbance," encouraged classification and removal from the mainstream of the child so labeled, and implied a genetic, or at least deep-rooted environmentally-based disorder. The new emphasis on teaching behaviors implies that student behavior can be effectively managed or directed in the current environment, thus encouraging

remediation. Such a shift in perspective can only be of benefit to children in both regular and special education.

References

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student. Alberta Journal of Educational Research, 1983, 29(2), 89-97.

Anderson, L. M., Evertson, C. M., & Brophy, J. E. An experimental study of effective teaching in first grade reading groups. Elementary School Journal, 1979, 79(4), 1, 193-223.

Aronfreed, J. Aversive control of socialization. In W. J. Arnold (Ed.), Nebraska Symposium on Motivation 1968. Lincoln: University of Nebraska Press, 1968.

Becker, W. C., Engelmann, S., Thomas, D. R. Teaching I: Classroom management. Chicago: Science Research Associates, 1975.

Borg, W. R., & Ascione, F. R. Classroom management in elementary mainstreaming classrooms. Journal of Educational Psychology, 1982, 74, 85-95.

Brophy, J. Teacher praise: A functional analysis. Review of Educational Research, 1981, 51(1), 5-32.

Brophy, J. E., & Evertson, C. M. Learning from teaching: A developmental perspective. Boston: Allyn & Bacon, 1976.

Cooper, M. L., Thomson, C. L., & Baer, D. M. The experimental modification of teacher attending behavior. Journal of Applied Behavior Analysis, 1970, 3, 153-157.

Conch, J. V., & Clement, T. H. Free time as a reinforcer for on-task behavior in a special education classroom. Psychological Reports, 1981, 48, 369-370.

Deno, S. L. Direct observation approach to measuring classroom behavior: Procedures and applications (Research Report No. 6). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979. (ERIC Document Reproduction Service No. ED 185 750)

Deur, J. L., & Parke, R. D. Effects of inconsistent punishment on aggression in children. Developmental Psychology, 1970, 2(5), 403-411.

Evertson, C. M., & Emmer, E. T. Effective management at the beginning of the school year in junior high classes. Journal of Educational Psychology, 1982, 74(4), 485-498.

Fisher, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. Teaching behaviors, academic learning time, and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, D. C.: National Institute of Education, 1980.

Gluek, S., & Gluek, E. Unraveling juvenile delinquency. Cambridge, MA: Harvard University Press, 1952.

Good, T. L., & Grouws, D. A. Teaching effects: A process-product study in fourth grade mathematics classes. Journal of Teacher Education, 1977, 28, 45-50.

Good, T., Grouws, D. A., & Beckerman, T. M. Curriculum pacing: Some empirical data in mathematics. Journal of Curriculum Studies, 1978, 10, 75-81.

Hall, R. V., Lund, D., & Jackson, D. Effects of teacher attention on study behavior. Journal of Applied Behavior Analysis, 1968, 1, 1-2.

Hansen, C. H., McLaughlin, T. F., Hansaker, D., & Young, J. W. A sports activity as a reinforcer for in-seat-behavior and generalization for academic responding in a special education student. Corrective and Social Psychiatry and Journal of Behavior Technology, Methods and Therapy, 1981, 27(3), 115-119.

Haring, N. G., & Phillips, E. L. Analysis and modification of classroom behavior. Englewood Cliffs, NJ: Prentice Hall, 1972.

Haynes, S. N. Principles of behavioral assessment. New York: Gardner Press, 1978.

Herman, S. I., & Tramontana, J. Instructions and group vs. individual reinforcement in modifying disruptive group behavior. Journal of Applied Behavior Analysis, 1971, 4, 113-119.

Hoge, R. D., & Luce, S. Predicting academic achievement from classroom behavior. Review of Educational Research, 1979, 49, 479-496.

Jones, F. H., & Miller, W. H. The effective use of negative attention for reducing group disruption in special elementary school classrooms. Psychological Record, 1974, 24, 435-448.

Kazdin, A. E. The token economy: A review and evaluation. New York: Plenum, 1977.

Kazdin, A. E., & Klock, J. The effect of nonverbal teacher approval on student attentive behavior. Journal of Applied Behavior Analysis, 1973, 6, 643-654.

Keller, H. R. Issues in the use of observational assessment. School Psychology Review, 1980, 9(1), 21-30.

Kounin, J. S. Discipline and group management in classrooms. New York: Holt, Rinehart & Winston, 1970.

Kratochwill, T. R., & Levin, J. R. What time-series designs may have to offer educational researchers. Contemporary Educational Psychology, 1978, 3, 273-329.

Lipinski, S., & Nelson, R. Problems in the use of naturalistic observation as a means of behavioral assessment. Behavior Therapy, 1974, 5, 341-351.

Madsen, C. H., Becker, W. C., & Thomas, D. R. Rules, praise, and ignoring: Elements of elementary classroom control. Journal of Applied Behavior Analysis, 1968, 1(2), 139-150.

Mattos, R. L., Mattson, R. H., Walker, H. M., & Buckley, N. K. Reinforcement and aversive control in the modification of behavior. Academic Therapy, 1969, 5(1), 37-52.

McGuire, W. H. Teacher burnout. Today's Education, 1979, 68, 5.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students. Journal of Special Education, in press.

National Institute of Education. Violent schools, safe schools. Washington, D.C.: U. S. Department of Health, Education and Welfare, 1978.

O'Leary, R. D., & Drabman, R. S. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

O'Leary, K. D., Kaufman, R. F., Kass, R. E., & Drabman, R. S. The effects of loud and soft reprimands on the behavior of disruptive students. Exceptional Children, 1970, 37, 145-155.

O'Leary, K. D., & Kent, R. N. Behavior modification for social action: Research tactics and problems. In L. A. Hamerlynck, L. C. Handy & E. J. Mash (Eds.), Behavior change: Methodology, concepts, and practice. Champaign, IL: Research Press, 1973.

O'Leary, K. D., & O'Leary, S. G. Classroom management: The successful use of behavior modification (2nd ed.). New York: Pergamon Press, 1977.

Pace, D. M., & Foreman, S. G. Variables related to the effectiveness of response cost. Psychology in the Schools, 1982, 19(3), 365-370.

Parke, R. D. Effectiveness of punishment as an interaction of intensity, timing, agent nurturance, and cognitive structuring. Child Development, 1969, 40, 213-235.

Parke, R. D., & Walters, R. H. Some factors influencing the efficacy of punishment training for inducing response inhibition. Monographs of the Society for Research in Child Development, 1967, 32 (No 1., Serial No. 109).

Powell, T. H., & Powell, I. Q. The use and abuse of using the timeout procedure for disruptive pupils. Pointer, 1982, 26, 18-22.

Robinson, P. W., Newby, T. J., & Ganzell, S. L. A token system for a class 1 for underachieving hyperactive children. Journal of Applied Behavior Analysis, 1981, 14(3), 307-315.

Rosenshine, B. Classroom instruction. In N. L. Gage (Ed.), The psychology of teaching methods. Chicago: University of Chicago Press, 1976.

Saal, F. E., Downey, R. G., & Lahey, M. A. Rating the ratings: Assessing the psychometric quality of rating data. Psychological Bulletin, 1980, 88(2), 413-428.

Sharpley, C. F., & Sharpley, A. M. Contingent vs. noncontingent rewards in the classroom: A review of the literature. Journal of School Psychology, 1981, 19(3), 250-259.

Skiba, R. J. Classroom behavior management: A review of the literature (Monograph No. 21). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1983.

Soli, S. D., & Devine, V. T. Behavioral correlates of achievement: A look at high and low achievers. Journal of Educational Psychology, 1976, 68(3), 335-341.

Stallings, J. Implementation and child effects of teaching practices in Follow Through classrooms. Monographs of the Society for Research in Child Development, 1975, 40, 7-8, Serial No. 163.

Stevens, R., & Rosenshine, B. Advances in research on teaching. Exceptional Education Quarterly, 1981, 2(1), 1-10.

Swanson, L. Removal of positive reinforcement to alter LD adolescents' preacademic problems. Psychology in the Schools, 1979, 16(2), 286-292.

Thompson, T., & Lubinski, D. On the kinetic structure of behavioral repertoires. Paper presented at the meeting of the Association of Behavior Analysts, Milwaukee, May, 1982.

Walters, R. H., Parke, R. D., & Cane, V. A. Timing of punishment and the observation of consequences to others as determinants of response inhibition. Journal of Experimental Child Psychology, 1965, 2, 10-30.

Footnote

This paper is based on the author's research toward completion of a master's thesis, which was supervised by Dr. James Terwilliger.

Table 1

Effective Teaching Behaviors: Parameters and Research Support

Variable	Effective Characteristics (Most Effective When:)	Cited In:
Positive Attention for Appropriate Behavior	- Contingent upon appropriate classroom behavior	Hall, Lund, & Jackson (1968); Madsen, Becker, Thomas (1968); Sharpley & Sharpley (1981)
	- Includes some non-verbal element	Kazdin & Klock (1973)
	- Specifies desired behavior	O'Leary & O'Leary (1979)
Back-up Reinforcers	- Tokens or points exchangeable for tangible reinforcer	Kazdin (1977)
	- Contingent upon desired classroom behavior	O'Leary & Drabman (1971)
Interventions for Inappropriate Behavior	- Delivered briefly and unemotionally	Jones & Miller (1974); O'Leary et al. (1970)
	- Delivered consistently across behaviors and students	Gluek & Gluek (1952); Deur & Parke (1970)
	- Specifies target behavior and/or desired alternative	Parke & Walters (1967); Anderson, Evertson & Brophy (1979)
Immediacy of Consequences	- Delivered immediately after onset of misbehavior	Aronfreed (1968); Walters, Parke, & Cane (1965); Evertson & Emmer (1982)
	- Delivered before misbehavior spreads	Jones & Miller (1974); Kounin (1970)
Classroom Rules	- Provide cues for students on where they should be and what they should be doing	Haring & Phillips (1972); Becker, Engelmann & Thomas (1975); Parke (1969)
	- Clear and specific	Evertson & Emmer (1982)
	- Applied consistently	

Table 1 (continued)

Variable	Effective Characteristics (Most Effective When:)	Cited In:
Teacher Awareness	- Teacher demonstrates awareness of student off-task or disruptive behaviors	Kounin (1970); Borg & Ascione (1982)
	- Teacher monitors student progress and completion of assignments	Evertson & Emmer (1982); Stallings, (1975)
Feedback & Corrections	- Teacher provides feedback on the correctness of responses on an individual basis	Brophy & Evertson (1976); Good & Grouws (1977); Anderson, Evertson, & Brophy (1979); Fisher et al. (1980)
Active Academic Responding	- More time allocated during school hours to academic subjects	Fisher et al. (1980); Brophy & Evertson (1976)
	- Lesson evokes active, observable responses	Stallings (1975)
Teacher Directed Learning	- The teacher, not the student is firmly in control of the lesson	Stevens & Rosenshine (1980); Fisher et al. (1980)
	- Instruction is focused on primarily academic topics	Stallings (1975); Rosenshine (1976)
Transitions	- Teacher provides cues before and during transitions that specify action students are to take	Kounin (1970); Anderson, Evertson, & Brophy (1979)
Lesson Pacing	- Lesson pace is rapid	Good et al. (1978); Stallings (1975);
	- Provides many opportunities for student response	Anderson, Evertson, & Brophy (1979)

Table 2

CMOS Scale Item Means

	TIME 1	TIME 2	TIME 3	OVERALL
	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)
SCALE 1:				
1. Feedback & Corrections	3.33 (1.63)	3.83 (1.60)	4.17 (1.33)	3.78 (.75)
2. Positive Attention	3.5 (1.05)	3.50 (1.51)	3.33 (1.37)	3.44 (.935)
Bonus: Back-up Reinforcers	1.33 (1.03)	1.00 (1.045)	1.37 (1.10)	1.11 (.54)
3. Interventions for Inappropriate Behavior	3.67 (1.75)	4.2 (.837)	3.83 (1.33)	3.94 (.98)
4. Immediacy of Consequences	3.83 (1.60)	2.8 (2.05)	3.67 (1.37)	3.53 (1.23)
5. Teacher Awareness	3.33 (1.51)	3.5 (1.38)	3.5 (.837)	3.44 (.96)
SCALE I $\bar{X}$	3.80 (1.36)	3.91 (1.33)	3.90 (1.05)	3.85 (.76)
SCALE 2:				
6. Classroom Rules	3.25 (2.06)	3.4 (1.67)	4.67 (.577)	3.57 (1.39)
7. Transitions	4.00 (1.41)	3.8 (1.64)	3.17 (1.72)	3.33 (1.32)
8. Lesson Structure	4.50 (.837)	4.33 (1.21)	4.0 (1.10)	4.23 (.65)
9. Active Acad Responding	4.0 (.894)	3.5 (1.52)	3.67 (1.21)	3.72 (1.02)
10. Pacing & Enthusiasm	3.67 (1.60)	3.33 (1.21)	3.17 (1.33)	3.39 (.77)
SCALE 2 $\bar{X}$	3.96 (.87)	3.72 (1.33)	3.59 (1.06)	3.69 (.93)
TOTAL SCALE $\bar{X}$	3.87 (1.11)	3.79 (1.28)	3.77 (.93)	3.77 (.84)

Table 3

Differences Between Primary (Gr 1-3) and Intermediate (Gr 4-6) Classrooms on CMOS Ratings

	Primary	Intermediate	
Teacher Variables			
Feedback and Corrections	3.33	4.22	
Positive Attention	3.22	3.67	
Back up Reinforcers	1.11	1.11	
Interventions for Inappropriate Behaviors	3.22	4.67	
Immediacy of Consequences	2.67	4.39	
Teacher Awareness	2.67	4.22	
Scale $\overline{X}$	3.24	4.46	
Classroom Rules	2.61	5.00 (n=2)	
Transitions	2.22	4.44	
Lesson Structure	3.78	4.78	
Active Academic Responding	2.89	4.56	
Pacing	2.89	3.89	
Scale $\overline{X}$	2.88	4.51	
Total $\overline{X}$	3.06	4.48	
Student Variables			Range of # of (Behaviors/Classroom)
Noise/Child	2.67	.28	0-32
Out of Seat/Child	1.64	.40	0-28
Off Task/Child	5.34	1.51	0-62
Total Behaviors per Child	9.67	2.13	0-122

Table 4

Mean Behaviors/Child

	Time 1	Time 2[a]	Time 3	Overall
	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)
Noise	1.11 (1.44)	.92 (1.65)	2.04 (3.04)	1.45 (1.95)
Out of Seat	.70 (.95)	.54 (.48)	1.64 (2.72)	1.02 (1.14)
Off Task	1.98 (1.70)	3.37 (2.33)	4.70 (5.55)	3.42 (2.74)
Total Behavior	3.83 (3.94)	4.83 (4.16)	8.38 (11.24)	5.90 (5.78)

[a]N=5; for Time 1 and Time 3, N=6.

Table 5

Percentage of Intervals in Which Behavior Occurred

	Time 1	Time 2	Time 3	Overall
	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)	$\bar{X}$ (SD)
Noise	.41 (.407)	.25 (.283)	.33 (.291)	.33 (.311)
Out of Seat	.31 (.298)	.19 (.120)	.28 (.333)	.26 (.173)
Off Task	.73 (.282)	.71 (.281)	.63 (.389)	.69 (.294)
Total Behavior	.48 (.296)	.47 (.260)	.42 (.317)	.46 (.257)

Table 6

Correlations Between Scale Ratings Aggregated Across Time

	Feedback and Corrections	Positive Attention	Back up Reinforcers	Interventions for Inappropriate Behaviors	Immediacy of Consequences	Teacher Awareness	Classroom Rules	Transitions	Lesson Structure	Active Academic Responding
Feedback and Corrections	1.000									
Positive Attention	.391	1.000								
Back up Reinforcers	.290	.320	1.000							
Interventions for Inappropriate Behaviors	.739	.105	-.363	1.000						
Immediacy of Consequences	.646	-.032	-.453	.982	1.000					
Teacher Awareness	.504	.305	-.369	.863	.853	1.000				
Classroom Rules	.119	.504	-.241	.230	.096	.268	1.000			
Transitions	.787	.415	.062	.836	.788	.880	.070	1.000		
Lesson Structure	.885	.195	.336	.733	.658	.585	.061	.834	1.000	
Active Academic Responding	.658	.365	.187	.673	.653	.810	-.112	.959	.780	1.000
Pacing and Enthusiasm	.793	.543	-.176	.860	.804	.830	.279	.874	.630	.728

Table 7

Correlations Among Student Behavior Codes[a]

	Noise	Out of Seat	Off-Task	Total
Noise	1.00			
Out of Seat	.892	1.00		
Off-Task	.875	.918	1.00	
Total	.947	.956	.972	1.00

[a]All correlations significant at $p \leq .01$.

Table 8

Correlations Between Teacher and Student Behaviors[a]

	Noise	Out of Seat	Off Task	Total Behaviors
Feedback and Correlations	.727	-.819	-.830	-.841
Positive Attention	.008	.056	-.289	-.138
Back-up Reinforcers	.134	.210	.246	.188
Interventions for Inappropriate Behavior	-.888*	-.913*	-.948*	-.970*
Immediacy of Consequences	-.835	-.884*	-.878	-.919*
Teacher Awareness	-.702	-.588	-.756	-.767
Classroom Rules[b]	-.960*	-.741	-.838	-.887
Transitions	-.726	-.661	-.772	-.813
Lesson Structure	-.860	-.768	-.742	-.839
Active Academic Responding	-.588	-.476	-.566	-.644
Pacing and Enthusiasm	-.628	-.714	-.884*	-.823
Scale Mean	-.831	-.735	-.828	-.879

[a]All correlations greater than ± .73 are significant at the .05 level.

[b]N=5, all other variables N=6

*$p \leq .01$

Appendix A

Classroom Management Observation Scale

Operational Definitions

I. TEACHER ATTENTION

1. Feedback and Corrections

5--Students receive feedback on the correctness of responses 80% or more of the time. For example, mispronunciations or skips are corrected by the teacher during oral reading, student work is monitored during independent seatwork, etc.

1--Feedback from the teacher is provided to less than 10% of student responses.

2. Positive Attention for Appropriate Behavior

5--Teacher praises student for appropriate classroom behavior or frequently follows desired behavior with some form of positive social attention, such as smiles, pats, physical proximity, etc. Such praise or attention occurs soon after the behavior it is in response to, and usually includes some specification of the desired behavior.

1--Students receive little or no attention or reward for appropriate classroom behavior. When praise is given, it is general in nature, or refers to a large block of time ("You're having a good day today, Susie"), and may even be sarcastic in nature.

Bonus--Add 2 rating points if present.

Back-up Reinforcers--Students receive points, free-time, food, privileges, or other tangible rewards contingent on desired behavior. Non-tangible reinforcers (points, stars, etc.) must be exchangeable for tangible reinforcers in order to earn the bonus.

3. Interventions for Inappropriate Behavior

5--Interventions (commands, reprimands or sharp glances, redirection of student behavior, detention, time-out, etc.) for student misbehavior are delivered by the teacher briefly, unemotionally, and consistently across behaviors and students. The targeted behavior is specified at the time of the intervention.

1--Teacher interventions for negative behavior are drawn-out (nagging, lecturing, etc.), critical, or delivered in a harsh or emotional manner. Reprimands are global ("your behavior is terrible") and/or refer to past as well as present behavior.

4. Immediacy of Consequences

5--Consequences for inappropriate classroom behavior--from reprimands to time out--are delivered immediately following the occurrence of the student behavior (within 5-10 seconds).

Operational Definitions (Cont'd)

1--Behavior is allowed to continue for some time, or even accelerate, before intervention takes place. Other students may be distracted from their work to observe the disruptive behavior, or join in disruptive behavior before teacher intervenes. Teacher may wait up to a minute or more to intervene, or may not intervene at all.

5. Teacher Awareness

5--When working with one student (one-on-one), the teacher is aware of other students in the classroom, giving attention to them (even at a distance), or specifying when they may receive such attention.

OR

When working with a group, the teacher is aware of, and strives to maintain the attention of, the entire group even when directing attention toward one student. All students receive an approximately equal proportion of teacher attention over the course of the observation period.

1--Students are ignored for long periods of time while the teacher works with one student or small group of students. One student may receive a disproportionate share of teacher attention over time, or the teacher may be engaged in an activity separate from students (such as paper work), allowing the students to work unmonitored.

II. CLASSROOM STRUCTURE

6. Classroom Rules

5--Students are provided with cues as to where they should be and what they should be doing, i.e., rules of the classroom are posted, directions are clear and specific, students are provided with clear reminders or consequences when not engaged in assigned activities. Rules and directions are applied consistently across students and behaviors.

1--Rules are not specified or appear not to be in effect. Student is free to pursue activities other than those prescribed by the teacher. When applied, rules or directions are inconsistent or "made up on the spur of the moment."

7. Transitions

5--Transitions are smooth and orderly. The teacher provides cues and directions before and during transition that specify the action students are to be taking ("Take out your math books and turn to page 58"). Transitions usually take less than two minutes.

1--Transitions between activities are spontaneous and disorganized, and may take up to five minutes or more. The teacher provides only very general cues and directions ("Now it's math time").

Operational Definitions (Cont'd)

8. Lesson Structure

5--The instruction received by the students is organized and businesslike; teacher is firm in direction and control of activities. Student is provided with questions, has material to cover, etc. Classroom discussion is limited to the subject being covered in the lesson, and rarely includes non-academic conversation.

1--The teacher's lesson is casually organized and very spontaneous. Teacher is not committed to having the student work on a particular set of materials, and the lesson may change depending on the mood of student or teacher. Instructional lessons may be interrupted to include "process" and other non-academic activities, such as discussing student or teacher interests or feelings at length.

9. Active Academic Responding

5--The curriculum or lesson evokes active, observable responses. The student is engaged in oral or written responding to teacher questions or written material, i.e., reading aloud, writing answers, responding to tape or teacher questions.

1--The curriculum involves primarily passive responding. The majority of student time is spent watching others, looking at instructional materials without making observable responses (includes silent reading), or listening to others read or recite.

10. Pacing and Enthusiasm

5--The pace of the lesson is rapid, with many opportunities for student response. The enthusiasm of the teacher is evident in frequent gestures, animated facial expressions, frequent attempts to involve students in discussion, etc.

1--The pace of the lesson is slow; the enthusiasm of students and/or teacher may be low. Few attempts are made to encourage student response.

Classroom Management Observation Scale

Teacher ID #_____ Date_____

Time Observation Begins:_____ Ends:_____

Number of Students in Class_____ Subject(Reading or Math):_____

I. TEACHER ATTENTION					
1. Feedback and Corrections	1	2	3	4	5
2. Positive Attention for Appropriate Behavior	1	2	3	4	5
Bonus: Backup Reinforcers? (+2)	☐				
3. Interventions for Inappropriate Behavior	1	2	3	4	5
4. Immediacy of Consequences	1	2	3	4	5
5. Teacher Awareness	1	2	3	4	5
II. CLASSROOM STRUCTURE					
6. Classroom Rules	1	2	3	4	5
7. Transitions	1	2	3	4	5
8. Lesson Structure	1	2	3	4	5
9. Active Academic Responding	1	2	3	4	5
10. Pacing and Enthusiasm	1	2	3	4	5

Appendix B

Classroom Behavior Frequency Count

Operational Definition

NOISE: Any verbalization or sound created by the child designed to disrupt or avoid the lesson as defined by the teacher. The noise may be generated vocally (includes "talkouts" or unintelligible noises) or non-vocally (as tapping a pencil or snapping fingers).

OUT OF PLACE: Any movement beyond either the explicitly or implicitly defined boundaries in which the child is allowed movement. If the child is seated at his desk, then movement of any sort out of the seat is "out of place." If the child is working with a group, then leaving the group is "out of place."

PHYSICAL CONTACT OR DESTRUCTION: Any contact with another person or another person's property which is unacceptable to that other person. Includes kicking, hitting, pushing, tearing, taking, breaking, etc., but may also include such neutral or positive contacts as hugging, if such contact is meant to dismay and is unacceptable to that person.

OFF TASK: Any movement off a prescribed activity which does not fall into one of the three previously defined categories, such as looking around, staring into space, observing inappropriate behavior on the part of other students, or doodling. A new behavior is counted any time the student engages in a different off task activity, even if he was previously off task. Neutral activities, such as looking up briefly, or scratching one's arm, are not counted if the student returns to the assigned task within 5 seconds.

B-2

Classroom Behavior Frequency Count

Teacher ID#____________________ Date____________

Number of Students in Classroom__________

Time observation Begins:________ Ends:________

NOISE	OUT OF PLACE	OFF TASK

Appendix C

Technical Characteristics of Measurement Instruments

Although observing and coding a larger sample of teacher behavior may produce more accurate descriptions of actual classroom conditions, increased complexity of observational instruments may lead to increases in measurement error (Lipinski & Nelson, 1974). Thus, this study also addressed the technical characteristics of the measurement instruments in an attempt to determine whether it is possible to reliably measure a number of teacher behaviors simultaneously. Valid statements concerning teacher and student behavior cannot be made unless reliability of the measurement system can be shown.

Technical Adequacy of the CMOS

In order to show evidence of validity, a measurement instrument must first evidence reliability--that is, freedom from measurement error. The reliability of observational assessment devices most often has been assessed through inter-observer agreement, internal consistency, and stability over time (Haynes, 1978). In addition, the special characteristics of rating scales necessitate that the researcher document the degree to which the scale is free from rating errors such as halo effect (Saal, Downey, & Lahey, 1980). A number of statistical methods have been suggested for this purpose and, where appropriate, have been applied to the Classroom Management Observation Scale (CMOS).

Stability over time and raters. Correlations between ratings that occurred during successive observations revealed that individual scales tended not to be stable over time. Of the 11 items, only

active academic responding evidenced even moderate correlations (.44-.65) between observations at Times 1, 2, and 3. The other variables fluctuated widely in their consistency over time: positive attention for appropriate behavior, for example, showed a correlation of -.06 between the first and second observations, .68 between the second and third observations, and .00 between the first and third observations. Results for the total scale score indicated that the entire scale was somewhat more stable over time, but still had a relatively low index of consistency across observations. For the total scale, correlations between Time 1 and Time 2, Time 2 and Time 3, and Time 1 and Time 3 were .29, .21, and .34, respectively. It is important to note that, given a different rater on each occasion, rater and observation occasion were confounded. Thus, this lack of stability could be attributed either to lack of observer agreement through observer drift or to wide variations in teacher behavior over observations. Either condition would be exacerbated by the small number of subjects.

Internal consistency. Mean inter-item correlations for each variable tended to be moderate for Times 1 and 3 and somewhat high for Time 2 (see Table C-1). Of the 11 items, only back-up reinforcers consistently failed to correlate with other variables, although this may be due in part to its measurement by means of a two-point scale, rather than five-point scale used for other items.

Insert Table C-1 about here

Halo and restriction of range effects. Inter-item correlations also provide a test for the presence of a halo effect, the tendency for a rater to assign similar ratings on a number of dimensions based on a global impression of the ratee (Saal et al., 1980). Higher inter-item correlations would tend to indicate the presence of halo effect. In addition, if a halo effect is present, one would expect averaging ratings over observers to result in lower inter-item correlations since this would tend to wash out individual rater error. For the CMOS, mean inter-item correlations tended to be moderate, and, with the exception of Time 2, inter-item correlations were higher when ratings were averaged across raters. Data collected during the second session revealed higher inter-item correlations; thus, halo effect may be evidenced at Time 2.

Restriction of range refers to the tendency of a rater to restrict ratings for all observations to a given range; it also has been referred to as central tendency, leniency, or severity, depending on the location of range restriction. The most common approach to assessing range restriction involves calculating the standard deviation of the ratings assigned to all ratees on a particular dimension: smaller standard deviations reflect greater range restriction (Saal et al., 1980). The standard deviations for the four observers in this study ranged from 0 to 2.65 (see Table C-2). Three of the four observers showed moderate to high variability in their ratings across teachers. Observer 3 showed fairly severe restriction of range on a number of scale items. Inspection of the means for these items revealed a uniform tendency for this observer to rate

teachers more highly than other observers, a possible sign of observer leniency.

Insert Table C-2 about here

Technical Characteristics of Behavior Counts

Stability over time and raters. The observations of student behavior produced data that were extremely stable over time and observer. High correlations between observations of total misbehaviors emitted during each observation session provide evidence of a high degree of stability between Times 1 and 2 (.97), and Times 2 and 3 (.93), and a moderate degree of consistency between Times 1 and 3 (.47). The inter-session correlations for specific behaviors (noise, out-of-seat, off-task) showed similar results. Thus, although rates of behavior emitted within classrooms tended to change over time (see below), relative rates of misbehavior between classrooms were remarkably stable.

Consistency across behavior codes and measures. A high degree of correlation also was evidenced among the specific behaviors observed. Correlations between proportion-of-interval and absolute frequency data were high, ranging from .86 for noises to .96 for off-task behaviors. Correlations among different behaviors remained high (.80 to .90), even under different statistical treatments.

Discussion

In the observation and measurement of classrooms, one must be careful to view a broad enough sample of the flux of events to ensure

"ecological validity" (Kratochwill & Levin, 1978). On the other hand, increasing the breadth and complexity of one's measuring instrument increases the risk of observer unreliability (Haynes, 1978).

The present study sought a middle route, observing a variety of behaviors simultaneously while defining those behaviors as operationally as possible, to reduce observer error. Yet limitations of the rating scale format, as well as time and resource constraints, may prevent firm conclusions about whether these goals were achieved. When ratings on a number of different scales prove highly intercorrelated, as they were for the second observation session, it is difficult, if not impossible, to detect whether these results follow from rater halo effect or the simultaneous presence of these behaviors in the teachers observed. Similarly, the failure of raters to achieve exact agreement, and the possible leniency errors of one of the raters provide qualifications to any conclusions based upon the results. But it is still impossible to determine the extent to which the extreme instability of ratings of teacher behavior over time was due to observer error, insufficient number of observations, or the situation-specificity of behavior. While increased time and resources for observer training and more extensive observations might allow for better experimental control, the unsolved methodological issues surrounding rating scales might set an upper limit on the returns that a better design might bring (cf. Saal et al., 1980).

In contrast to teacher behaviors, student behavior coding proved extremely stable over time. Although this might argue for the superiority of low inference coding systems, it also poses the

intriguing possibility that teachers cause stable effects while varying in specific behaviors. Some support for this hypothesis might be found in the fact that, although changing in frequency within classrooms, the relative degree of student misbehavior among classrooms was very stable. In fact, results indicated that while student misbehavior remained stable or decreased in classrooms with high teacher management ratings, classrooms with low ratings tended to show increases in misbehavior over time. Thus, although specific teacher behaviors may be variable from point to point in time, "mean" levels of those behaviors may serve to establish a "classroom climate" that will determine the overall effectiveness of classroom management.

Table C-1

Mean Inter-Item Correlations for the CMOS

	Time 1	Time 2	Time 3	Aggregate[a]
Feedback and Corrections	.628	.715	.418	.635
Positive Attention	.615	.402	.298	.294
Back-up Reinforcers	.257	.428	-.076	.272
Interventions for Inappropriate Behavior	.571	.493	.588	.626
Immediacy of Consequences	.620	.505	.456	.566
Teacher Awareness	.322	.629	.456	.605
Classroom Rules	.501	.747	.266	.622
Transitions	.544	.822	.563	.723
Lesson Structure	.625	.731	.257	.652
Active Academic Responding	-.111	.733	.431	.651
Pacing and Enthusiasm	.308	.731	.402	.656
Mean Inter-Item Correlation	.408	.634	.338	.455

[a]Represents the mean inter-item correlation when ratings are aggregated (by teacher) across rater and time.

Table C-2

Mean and Standard Deviation by Rater[a]

	Rater 1		2		3		4	
	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)	$\bar{X}$	(SD)
Feedback and Corrections	3.50	(1.29)	2.40	(1.34)	5.00	(0)	4.00	(1.00)
Positive Attention for Appropriate Behavior	3.75	(.96)	2.60	(1.14)	4.25	(.96)	3.33	(1.53)
Interventions for Inappropriate Behavior	3.50	(2.08)	3.60	(1.52)	5.00	(0)	3.33	(.33)
Immediacy of Consequences	3.00	(2.45)	3.40	(1.52)	5.00	(0)	2.33	(1.53)
Teacher Awareness	3.50	(1.92)	3.40	(1.14)	3.75	(.16)	2.67	(.58)
Classroom Rules	4.00	(1.83)	3.80	(2.59)	5.00	(0)	4.00	(1.00)
Transition	4.25	(1.71)	3.20	(1.64)	4.67	(.82)	3.00	(2.65)
Lesson Structure	4.50	(.58)	3.60	(1.34)	4.5	(1.00)	4.33	(1.16)
Active Academic Responding	4.00	(.41)	3.00	(1.58)	4.00	(1.16)	3.33	(.58)
Pacing and Enthusiasm	3.25	(.96)	2.60	(1.52)	4.00	(.82)	3.67	(.58)

[a]Ratings presented here are somewhat higher than actual means and may actually exceed 5, due to the coding of "not scorable" as a 6 in the calculation of standard deviations.

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $4.00 each, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid. Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

The publications listed here are only those that have been prepared since 1982. For a complete, annotated list of all IRLD publications, write to the Editor.

Wesson, C., Mirkin, P., & Deno, S. Teachers' use of self instructional materials for learning procedures for developing and monitoring progress on IEP goals (Research Report No. 63). January, 1982.

Fuchs, L., Wesson, C., Tindal, G., Mirkin, P., & Deno, S. Instructional changes, student performance, and teacher preferences: The effects of specific measurement and evaluation procedures (Research Report No. 64). January, 1982.

Potter, M., & Mirkin, P. Instructional planning and implementation practices of elementary and secondary resource room teachers: Is there a difference? (Research Report No. 65). January, 1982.

Thurlow, M. L., & Ysseldyke, J. E. Teachers' beliefs about LD students (Research Report No. 66). January, 1982.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). January, 1982.

King, R., Wesson, C., & Deno, S. Direct and frequent measurement of student performance: Does it take too much time? (Research Report No. 67). February, 1982.

Greener, J. W., & Thurlow, M. L. Teacher opinions about professional education training programs (Research Report No. 68). March, 1982.

Algozzine, B., & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept (Research Report No. 69). March, 1982.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions (Research Report No. 70). March, 1982.

Shinn, M. R., Ysseldyke, J., Deno, S., & Tindal, G. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving (Research Report No. 71). March, 1982.

Thurlow, M. L. Graden, J., Greener, J. W., & Ysseldyke, J. E. Academic responding time for LD and non-LD students (Research Report No. 72). April, 1982.

Graden, J., Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student (Research Report No. 74). April, 1982.

Wesson, C., Deno, S., & Mirkin, P. Research on developing and monitoring progress on IEP goals: Current findings and implications for practice (Monograph No. 18). April, 1982.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students (Research Report No. 75). May, 1982.

Algozzine, B., Ysseldyke, J., Christenson, S., & Thurlow, M. Teachers' intervention choices for children exhibiting different behaviors in school (Research Report No. 76). June, 1982.

Tucker, J., Stevens, L. J., & Ysseldyke, J. E. Learning disabilities: The experts speak out (Research Report No. 77). June, 1982.

Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). July, 1982.

Mirkin, P. K., & Potter, M. L. A survey of program planning and implementation practices of LD teachers (Research Report No. 80). July, 1982.

Fuchs, L. S., Fuchs, D., & Warren, L. M. Special education practice in evaluating student progress toward goals (Research Report No. 81). July, 1982.

Kuehnle, K., Deno, S. L., & Mirkin, P. K. Behavioral measurement of social adjustment: What behaviors? What setting? (Research Report No. 82). July, 1982.

Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). July, 1982.

Videen, J., Deno, S., & Marston, D. Correct word sequences: A valid indicator of proficiency in written expression (Research Report No. 84). July, 1982.

Potter, M. L. Application of a decision theory model to eligibility and classification decisions in special education (Research Report No. 85). July, 1982.

Greener, J. E., Thurlow, M. L., Graden, J. L., & Ysseldyke, J. E. The educational environment and students' responding times as a function of students' teacher-perceived academic competence (Research Report No. 86). August, 1982.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). August, 1982.

Skiba, R., Wesson, C., & Deno, S. L. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 88). September, 1982.

Marston, D., Tindal, G., & Deno, S. L. Eligibility for learning disability services: A direct and repeated measurement approach (Research Report No. 89). September, 1982.

Thurlow, M. L., Ysseldyke, J. E., & Graden, J. L. LD students' active academic responding in regular and resource classrooms (Research Report No. 90). September, 1982.

Ysseldyke, J. E., Christenson, S., Pianta, R., Thurlow, M. L., & Algozzine, B. An analysis of current practice in referring students for psychoeducational evaluation: Implications for change (Research Report No. 91). October, 1982.

Ysseldyke, J. E., Algozzine, B., & Epps, S. A logical and empirical analysis of current practices in classifying students as handicapped (Research Report No. 92). October, 1982.

Tindal, G., Marston, D., Deno, S. L., & Germann, G. Curriculum differences in direct repeated measures of reading (Research Report No. 93). October, 1982.

Fuchs, L.S., Deno, S. L., & Marston, D. Use of aggregation to improve the reliability of simple direct measures of academic performance (Research Report No. 94). October, 1982.

Ysseldyke, J. E., Thurlow, M. L., Mecklenburg, C., & Graden, J. Observed changes in instruction and student responding as a function of referral and special education placement (Research Report No. 95). October, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study (Research Report No. 96). November, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Report No. 97). November, 1982.

Tindal, G., Wesson, C., Germann, G., Deno, S. L., & Mirkin, P. K. The Pine County model for special education delivery: A data-based system (Monograph No. 19). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. An analysis of the conceptual framework underlying definitions of learning disabilities (Research Report No. 98). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. Public-policy implications of different definitions of learning disabilities (Research Report No. 99). November, 1982.

Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Wesson, C., Deno, S. L., & Algozzine, B. Generalizations from five years of research on assessment and decision making (Research Report No. 100). November, 1982.

Marston, D., & Deno, S. L. Measuring academic progress of students with learning difficulties: A comparison of the semi-logarithmic chart and equal interval graph paper (Research Report No. 101). November, 1982.

Beattie, S., Grise, P., & Algozzine, B. Effects of test modifications on minimum competency test performance of third grade learning disabled students (Research Report No. 102). December, 1982

Algozzine, B., Ysseldyke, J. E., & Christenson, S. An analysis of the incidence of special class placement: The masses are burgeoning (Research Report No. 103). December, 1982.

Marston, D., Tindal, G., & Deno, S. L. Predictive efficiency of direct, repeated measurement: An analysis of cost and accuracy in classification (Research Report No. 104). December, 1982.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. Teaching structure and student achievement effects of curriculum-based measurement: A causal (structural) analysis (Research Report No. 105). December, 1982.

Mirkin, P. K., Fuchs, L. S., & Deno, S. L. (Eds.). Considerations for designing a continuous evaluation system: An integrative review (Monograph No. 20). December, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). December, 1982.

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). January, 1983.

Thurlow, M. L., Ysseldyke, J. E., & Casey, A. Criteria for identifying LD students: Definitional problems exemplified (Research Report No. 108). January, 1983.

Tindal, G., Marston, D., & Deno, S. L. The reliability of direct and repeated measurement (Research Report No. 109). February, 1983.

Fuchs, D., Fuchs, L. S., Dailey, A. M., & Power, M. H. Effects of pre-test contact with experienced and inexperienced examiners on handicapped children's performance (Research Report No. 110). February, 1983

King, R. P., Deno, S., Mirkin, P., & Wesson, C. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 111). February, 1983.

Tindal, G., Deno, S. L., & Ysseldyke, J. E. Visual analysis of time series data: Factors of influence and level of reliability (Research Report No. 112). March, 1983.

Tindal, G, Shinn, M., Fuchs, L., Fuchs, D., Deno, S., & Germann, G. The technical adequacy of a basal reading series mastery test (Research Report No. 113). April, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Communication of IEP goals and student progress among parents, regular classroom teachers, and administrators using systematic formative evaluation (Research Report No. 114). April, 1983.

Wesson, C. Two student self-management techniques applied to data-based program modification (Research Report No. 115). April, 1983.

Wesson, C., Skiba, R., Sevcik, B., King, R., Tindal, G., Mirkin, P., & Deno, S. The impact of the structure of instruction and the use of technically adequate instructional data on reading improvement (Research Report No. 116). May, 1983.

Wesson, C. Teacher vs student selection of instructional activities (Research Report No. 117). May, 1983.

Tindal, G., & Deno, S. Factors influencing the agreement between visual and statistical analyses of time series data (Research Report No. 118). June, 1983.

Skiba, R. S. Classroom behavior management: A review of the literature (Monograph No. 21), June, 1983.

Graden, J. L., Thurlow, M. L., & Ysseldyke, J. E. When are students most academically engaged? Academic responding time in different instructional ecologies (Research Report No. 119). June, 1983.

Fuchs, L. S., Deno, S. L., & Roettger, A. The effect of alternative data-utilization rules on spelling achievement: An n of 1 study (Research Report No. 120). June, 1983.

Skiba, R., Sevcik, B., Wesson, C., King, R., & Deno, S. The non-effect of process-product variables in resource classrooms (Research Report No. 121). June, 1983.

Fuchs, L. Tindal, G., Shinn, M., Fuchs, D., Deno, S., & Germann, G. Technical adequacy of basal readers' mastery tests: The Ginn 720 series (Research Report No. 122). June, 1983.

Tindal, G., Germann, G., Marston, D., & Deno, S. The effectiveness of special education: A direct measurement approach (Research Report No. 123). June, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Curriculum-based measurement: Effects on instruction, teacher estimates of student progress, and student knowledge of performance (Research Report No. 124). July, 1983.

Skiba, R., Marston, D., Wesson, C., Sevcik, B., & Deno, S. L. Characteristics of the time-series data collected through curriculum-based reading measurement (Research Report No. 125). July, 1983.

Ysseldyke, J., Christenson, S., Graden, J., & Hill, D. Practical implications of research on referral and opportunity to learn (Monograph No. 22). July, 1983.

Marston, D., Deno, S., & Tindal, G. A comparison of standardized achievement tests and direct measurement techniques in measuring pupil progress (Research Report No. 126). July, 1983.

Fuchs, D., Fuchs, L. S., Tindal, G., & Deno, S. L. Variability of performance: A "signature" characteristic of learning disabled children? (Research Report No. 127). July, 1983.

Tindal, G., Fuchs, L., Fuchs, D., Shinn, M., Deno, S., & Germann, G. The technical adequacy of a basal series mastery test: The Scott-Foresman reading program (Research Report No. 128). July, 1983.

Fuchs, L. S., Fuchs, D., & Deno, S. L. The nature of inaccuracy among readability formulas (Research Report No. 129). July, 1983.

Fuchs, L., Tindal, G., Fuchs, D., Shinn, M., Deno, S., & Germann, G. The technical adequacy of a basal reading mastery test: The Holt basic reading series (Research Report No. 130). July, 1983.

Ysseldyke, J. E., Christenson, S., Algozzine, B., & Thurlow, M. L. Classroom teachers' attributions for students exhibiting different behaviors (Research Report No. 131). July, 1983.

Tindal, G., Germann, G., & Deno, S. Descriptive research on the Pine County norms: A compilation of findings (Research Report No. 132). July, 1983.

Skiba, R. J. The relationship between classroom management strategies and student misbehaviors (Research Report No. 133). July, 1983.